

Overdetermination and Elimination

I focus on two arguments, due to Jaegwon Kim and Trenton Merricks, that move from claims about the sufficiency of one class of causes to the reduction or elimination of another class of entity, via claims about overdetermination. I argue that in order to validate their move from sufficiency to reduction or elimination, both Kim and Merricks must assume that there can be no 'weak overdetermination'; i.e. that no single effect can have numerically distinct but dependently sufficient causes occurring at the same time. One problem for both arguments is that weak overdetermination isn't obviously objectionable. That point has been well made before. But I want here to go further than merely shifting the burden of proof onto the advocates of overdetermination arguments. I want to tease out why they are so convinced that we must resist weak overdetermination and explain why their conviction is misguided. Both Merricks and Kim, I shall argue, ultimately rest their case on the same motivating principle, which I call the principle of additional causal powers. This principle, I argue further, should be rejected. It lacks argumentative support, and it begs the question against those at whom the arguments are directed.

1. Introduction

The focus of this paper is on two arguments that move from claims about the sufficiency of one class of entity to the reduction or elimination of another class of entity, via claims about overdetermination. The first argument is Jaegwon Kim's influential exclusion argument from the philosophy of mind (see e.g. Kim 1998, 2005). The second argument is Trenton Merricks's (2001) argument for eliminativism about ordinary objects. In brief, Kim argues that given that physical properties are causally sufficient for any would-be-effects of mental properties, and given that there is no widespread causal overdetermination, we must either identify mental properties with physical properties or accept that they are causally inert. Similarly, Merricks argues that given that particles arranged object-wise are causally sufficient for any would-be-effect of objects, and given that there is no widespread causal overdetermination, we must either identify objects with the physical particles that make them up or accept that they are causally inert.¹ Merricks argues further that if there *were* any ordinary objects they *would* do causal work, and that ordinary objects *could not* be identified with the particles that make them up at any given time. He concludes that there are no ordinary objects. Although Kim doesn't normally present his argument as eliminativist, it too has fairly immediate eliminativist consequences. First, if we assume with Kim that, as far as properties are concerned, to be is to have causal powers, the choice Kim offers is really between identification and elimination. Second, if we assume that mental properties are multiply realizable by physical properties, then it seems that mental properties *could not* be identified with physical properties.² For, whatever else identity might require, it requires at least coextensionality. These two assumptions are commonplace. If they are right, the real moral of Kim's argument is that there are no mental properties.

Kim's argument has inspired a huge debate, and has prompted various significant ontological and metaphysical revisions. Merricks's argument has caused less of a stir, but those who find Kim's argument compelling ought also to find Merricks's argument compelling. For the arguments share the same core structure. We can formulate this as follows:

¹ This argument is also hinted at, but not developed, by Peter van Inwagen (1990: 118, 122).

² See e.g. Putnam (1967), Fodor (1974).

1. Xs are causally sufficient for the would-be effects of Ys
 2. The would-be effects of Ys are not causally overdetermined
- Therefore,
- C. Ys are either identical with Xs or they are causally inert

If we replace 'X' with 'physical causes' and 'Y' with 'mental causes' we get the core of Kim's argument against non-reductive physicalism. If, on the other hand, we replace 'X' with 'particles arranged objectwise' and 'Y' with 'material object' we get the core of Merricks's argument against objects. Both arguments employ a number of further claims, but unless this core structure is sound, neither argument gets off the ground.³ I shall argue that this core is rotten.

The core argument is not valid as it stands and Kim and Merricks must rely on extra claims to validate their versions of it. To this end, Merricks introduces a much weaker than ordinary notion of overdetermination, while Kim puts down an extra principle – the exclusion principle – designed to take us from the denial of ordinary overdetermination to identity. I will argue that although these strategies look different on their face, they boil down to the same thing; namely, to denying that weak overdetermination of the Merricksian kind ever occurs. One problem for both arguments is that it is far from obvious why we ought to find weak overdetermination objectionable; a point well made by Theodore Sider (2003). But I want here to go further than merely shifting the burden of proof onto the advocates of overdetermination arguments. I want to tease out *why* they are so convinced that we must resist this kind of overdetermination and explain why their conviction is misguided. Both Merricks and Kim, I shall argue, ultimately rest their case on the same motivating principle, which I shall call the principle of additional causal powers. This principle, I argue further, is untenable. It lacks argumentative support, and it begs the question against those at whom the arguments are directed.

2. The Sufficiency Claims

The principal focus of this paper is the move from the causal sufficiency of one entity to the reduction or elimination of another. But before turning to this, I want to say a few words about the sufficiency claims that Kim and Merricks's arguments employ and why I think we should grant them.

It seems to me that Merricks is right when he claims that particles arranged objectwise are causally sufficient for the would-be effects of objects. For instance, when some particles arranged baseballwise are hurled against a windowpane, to use Merricks's example, these particles are sufficient, in the circumstances, for the window shattering. Once you have these particles acting in concert, nothing further needs to be added in order to bring about the effect. The circumstances are very important, of course, but I take it that things are only ever sufficient given the circumstances, and that any claim about sufficiency must be taken to implicitly assume this. What's more, Merricks's sufficiency claim concerns not just particles, but particles *arranged objectwise*. Given what Merricks means by this, his sufficiency claim is practically irresistible.⁴ Here is his definition of what it is for particles, or atoms as he calls them, to be arranged statuewise:

³ Indeed, this is true not just of Kim and Merricks's arguments, but of overdetermination arguments more generally. What we have here is the core not just of the principal argument *against* non-reductive physicalism, but also of the principal argument *in favour* of physicalism (see e.g. Papineau 2002). My discussion will apply to these arguments too.

⁴ I should say that Merricks talks of particles being arranged as particular *types* of objects (e.g. 'baseballwise', 'tablewise', 'chairwise'), but I shall use the term 'objectwise' to keep things simple. Those who worry that there is no such thing as an objectwise arrangement of particles, on grounds that 'object' is not a sortal term, can take 'objectwise' to be a placeholder for more specific terms.

Atoms are *arranged statuenwise* if and only if they both have the properties and also stand in the relations to microscopica upon which, if statues existed, those atom's *composing a statue* would non-trivially supervene. (Merricks 2001: 4)

Assuming that this definition applies to objects more generally, the circumstances where particles are arranged objectwise are precisely the ones in which the particles compose objects if there are any objects. More, the would-be effects of objects are those that objects would cause if they existed. So it is hard to see how particles arranged object-wise could fail to be causally sufficient for any would-be effect of objects. It is important to note also that this sufficiency cannot it seems be explained by particles' arranged objectwise being *identical* to objects. For objects have properties that particles arranged objectwise do not, such as being able to survive changes in their parts. But then objects could not be identical to the particles that compose them at any given time.

Kim's claim, that physical causes are sufficient for the would-be effects of mental causes, is perhaps not as immediately compelling as Merricks's claim. But, at least as far as the would-be *physical* effects of mental causes are concerned, it falls out of a thesis that is accepted by most contemporary philosophers of mind. This thesis is known as the causal closure of the physical, or the completeness of physics, and states that "physics is causally and explanatorily *self-sufficient*: there is no need to go outside the physical domain to find a cause, or a causal explanation, of a physical event." (Kim 2005: 16). If true, then any would-be physical effect of a mental cause has sufficient physical causes, such that nothing outside the physical domain is ever required to explain their occurrence. More, given another widely accepted thesis, the thesis of the supervenience of the mental on the physical, it seems that the mental must have some physical effects if it is to have any effects at all. There are various ways in which we might characterize supervenience, but the core idea common to all formulations is that if A supervenes on B, then no change in A is possible without some corresponding change in B. So, if the mental supervenes on the physical, then no change is possible in the mental without a corresponding change in the physical. And that, it is reasonable to think, means that mental causes must cause some physical effects in order to cause anything at all.⁵

3. Standard Overdetermination, Weak Overdetermination and the Exclusion Principle

Let us turn then to the move from sufficiency to identification or elimination. The first thing to note is that of course a given effect's having a sufficient cause X does not entail that the effect has no other sufficient cause distinct from X. It is not only that an effect may have, and almost certainly will have, a number of sufficient causes occurring at different times. It may also have distinct sufficient causes occurring at the very same time.⁶ Given this, the move from the sufficiency of one cause to the elimination of another relies crucially on the denial of overdetermination. Whether the denial of overdetermination warrants this move, however, depends very much on the notion of overdetermination in play.

The cases most naturally thought of as ones of overdetermination are cases where two causes operate through two distinct causal chains or nexuses; for example, when a window breaks from the impact of two baseballs thrown at it at once, or when an unfortunate man has a fatal heart attack at the same time as (but

⁵ Strictly speaking, this doesn't follow. It would be sufficient, for instance, if something *other* than the mental cause simultaneously caused the required physical change, but I ignore this possibility here.

⁶ As is well known, the possibility of multiple sufficient causes raises some problems for counterfactual analyses of causation. But this is generally thought to show that something is amiss with counterfactual analyses rather than that we are wrong in thinking that effects might have multiple causes.

not because of) being struck by lightning. What is distinctive about such cases is not just that there are *distinct* causes, but that there are *distinct and independent* causes. The independence here consists in the fact that had either cause been removed from the circumstances (and not replaced by anything) the effect would have been brought about by the other cause. The death of the unfortunate man is overdetermined in this way, for instance, because had lightning not struck him, the heart attack would still have killed him, and had he not had the heart attack he would still have been killed by the lightning. Overdetermination of this sort is well named. For in all such cases there is a genuine *excess* of causes for the same effect, fitting the ‘over’ in overdetermination. Indeed, it is tempting to reserve the term ‘overdetermination’ for overdetermination of this kind. For the purposes of the ensuing discussion, however, I will refer to this as ‘standard overdetermination’.

On the standard reading of overdetermination, everyone should be happy to grant the second premise of the core argument, whether we plug in Kim’s or Merricks’s values for the variables. For whether or not we can take objects and their composing particles, or mental causes and their accompanying physical causes, to be *distinct* causes, we surely cannot take them to be *distinct and independent* causes. Indeed, I know of nobody who wants to say that mental events cause things independently of happenings in the brain or that objects cause things independently of the matter out of which they are constituted. On this reading of the second premise, however, the argument is not valid. The causal sufficiency of Xs together with the denial of standard overdetermination could establish only the conclusion that the Ys do not cause their effects *independently of* Xs and that is considerably weaker than the intended conclusion. This is very important. It means that you cannot get from the conclusion that entities of a given kind do not cause things *independently of* entities of a second kind to the conclusion that entities of the first kind must be *identified* with entities of the second kind in order to cause anything at all without employing some *further* assumption.

Merricks bridges this gap by employing a much weaker notion of overdetermination than the standard one, offering the following sufficient condition:

An effect is overdetermined if the following are true: that effect is caused by an object; that object is causally irrelevant to whether some other—i.e. numerically distinct—object or objects cause that effect; and the other object or objects do indeed cause that effect. (2001: 58)⁷

Causal irrelevance, in turn, is defined as follows: “O is causally irrelevant to whether the xs [i.e. O’s constituent particles], acting in concert, cause a certain effect E” if the following four conditions are met (i) O is not identical to one of the xs, (ii) O is not a partial cause of E alongside the xs, (iii) none of the xs cause O to cause E, and (iv) O does not cause any of the xs to cause E (2001: 58). The term ‘causal irrelevance’ is at best misleading, given this definition.⁸ But we can reformulate Merricks’s sufficient

⁷ Merricks here intends ‘other objects’ to include the particles composing ordinary objects (or, rather, the particles that would compose ordinary objects if there were any).

⁸ Just consider how the notion plays out in the case of objects and their constituent particles. Given this construal of causal irrelevance, we must grant that every object is causally irrelevant to its constituent atoms’ causing their effects. Objects are not identical to any of the particles that compose them. Nor can objects and their particles plausibly be seen as either partial joint causes or distinct links in a causal chain. In granting this, however, we ought to be very clear on just what the causal irrelevancy claim amounts to. It says that objects are not relevant *in the ways specified in (i) to (iv)* to their parts causing their effects. It is perfectly compatible with this for objects to *cause the same effects* as their parts do. And assuming that anything that causes an effect is causally relevant to that effect, objects may be causally *relevant* to their parts’ effects, even though they are causally *irrelevant* (in the sense of i-iv) to their parts *causing* these effects. What’s more, conditions (i)-(iv) do not even rule out objects’ being causally relevant to their parts causing their effects in some other, as yet unspecified way. It seems perfectly compatible with the claim that objects are causally

condition in neutral terms. Where e is the effect and $c1$ and $c2$ are the purported overdetermining causes, Merricks sufficient condition for overdetermination breaks down as follows:

- 1) e is caused by $c1$
- 2) e is caused by $c2$
- 3) $c1$ is not identical to $c2$
- 4) $c1$ is not a partial cause of e alongside $c2$
- 5) $c1$ does not cause $c2$ to cause e
- 6) $c2$ does not cause $c1$ to cause e

Pulling this together, an effect is overdetermined on Merricks's account if it has causes that are not identical (3), are not merely partial causes, but are each sufficient for the effect in question (1, 2, 4), and are not parts of a causal chain (5, 6). In brief:

MOD: An effect is overdetermined if it has numerically distinct sufficient causes that are not distinct links of a causal chain.⁹

I shall refer to cases where this sufficient condition is met, but where the causes in question are not independent in the way characteristic of standard overdetermination, as cases of *weak* overdetermination, or *Merricksian* overdetermination.

If we adopt Merricks's sufficient condition for overdetermination, it is clear that any effect of an ordinary object would be overdetermined.¹⁰ More, if any effect of an ordinary object would be overdetermined, then of course any effect that is *not* overdetermined is not caused by an ordinary object. So on this reading, Merricks's argument is valid. As we shall see, however, it is far from clear that we have any good reason to deny that ordinary objects and their parts overdetermine their effects in this sense. But before I go on to argue this, let us look at the other way Merricks might have gone to validate his argument; Kim's way.

Kim proposes to bridge the gap between sufficiency and identity with his exclusion principle:

Exclusion Principle: "[n]o single event can have more than one sufficient cause occurring at any given time—unless it is a genuine case of overdetermination." (Kim, 2005, p.42)

By 'genuine overdetermination' Kim is explicit that he means *standard* overdetermination in the sense clarified above; i.e. overdetermination that "involves two or more separate and independent causal chains intersecting at a common effect." (Kim 2005: 48). Further, by 'more than one' Kim *has to* mean 'numerically distinct'. Otherwise the principle could not establish the *identity* of mental and physical causes, and without this his argument will not succeed. So Kim's exclusion principle must be understood as saying that no single event can have numerically distinct sufficient causes occurring at any given time—unless the causes in question are independent. Which entails that:

EP*: no single effect can have numerically distinct and *dependent sufficient* causes occurring at the same time.

irrelevant in this sense, for instance, that the parts of an object cause their effects *precisely because* they compose objects that instantiate certain properties. It is in light of this that I think the notion of causal irrelevance is best dropped out of the discussion.

⁹ My formulation here is more general than Merricks's formulation, as it is not restricted to objects, but I cannot see any reason why he would object to it.

¹⁰ This is assuming, as granted above, that particles arranged objectwise are causally sufficient for any such effect, and that particles arranged objectwise are numerically distinct from any objects they might compose.

That, in turn, is to say that there cannot be *weak* overdetermination of the Merricksian variety. So although Kim and Merricks's strategies look different on their face, the key sentiment is the same. This has important implications. First, if there was strong support for the exclusion principle that would much strengthen Merricks's argument. Moreover, it would pave the way for the use of overdetermination considerations in arguing for a more widespread elimination of entities of our ordinary ontology. If, on the other hand, there are good reasons to think both that baseballs and that particles arranged baseballwise are causally sufficient for window breakings, then the exclusion principle has to be rejected and Kim's argument fails.

4. *Sufficiency and Redundancy*

The million dollar question is this: Why exactly do Kim and Merricks think that effects cannot have distinct but dependent sufficient causes? Surprisingly, one finds very little in the way of an answer to this question in either Kim or Merricks. Kim claims that the exclusion principle is "practically an analytic truth with not much content" (2005: 51) and Merricks apparently feels the same way about his claim that "we should oppose systematic overdetermination on its own demerits." (2001: 147). But if Kim and Merricks think that the denial of weak overdetermination is something that we must accept without argument, they are wrong.

First, let us be clear that Kim's exclusion argument is *not* 'practically an analytic truth, with not much content'. To be sure, there is a principle in the vicinity of the exclusion principle that fits this description. Namely, that no single event can have more than one *independent* sufficient cause occurring at any given time—unless it is a genuine case of overdetermination. But that is not Kim's exclusion principle, and it is difficult to see how *that* principle could be any use in validating his argument. The important content of the exclusion principle, the part that takes us from the denial of overdetermination to the *identification* or elimination of causes, is the claim that no single effect can have numerically distinct and *dependent sufficient* causes occurring at the same time. That is to claim that there can be no weak overdetermination and that, I submit, is a very contentious claim.

Of course, if Merricks is right, then weak overdetermination ought to be opposed on its own demerits. But although Merricks evidently does not think this claim requires clarification, it is far from obvious what these demerits might be. This is a point well made by Sider (2003). It would be very mysterious, of course, if *standard overdetermination* were systematic and pervasive, as the autonomy of the causes in question would make this an astonishing *coincidence* at best. But, as Sider points out, such worries simply do not apply where the causes in question are as intimately related as objects and their parts. Widespread overdetermination of that kind would be readily explained by the intimate relationship holding between the causes and so would be neither mysterious nor coincidental. Nor does there seem to be any metaphysical incoherence involved in thinking that there might be widespread overdetermination of this sort; which is the other obvious objection that springs to mind. If there were such incoherence, one would expect it to stem from some form of incompatibility with the nature of causation. But there does not appear to be any such

incompatibility.¹¹ These points seem to me exactly right and I have nothing much to add to them. I want instead to press the question of what *else* might be wrong with weak overdetermination.

Merricks discusses a number of considerations in his book *Objects and Persons* that concern whether overdetermination is always objectionable, but mostly his discussion is defensive and consists in rebutting arguments that he imagines people might have for why overdetermination is *not* always objectionable.¹² As I find the considerations he offers on behalf of his imagined objectors largely besides the point I shall not discuss them here. I will focus instead on a positive reason that Merricks gives for why we *ought* always reject overdetermination. Namely, that overdetermination always involves *redundant* causes and redundancy is intrinsically bad. I think this worry is implicit in much of what Merricks says, but the clearest statement of it is probably found in the following passage:

Consider a substance dualist ... who, conceding causal closure of the physical, says that mental events cause physical events only by overdetermining the effects of physical causes. Pre-theoretically, that's an ugly picture. The redundancy is *all by itself* a reason to resist this form of substance dualism. More generally, we always have a reason to resist systematic causal overdetermination, along with any view that implies it. (66-67, my underlining, original italics).

As far as I can see, this is the only positive reason that Merricks gives us for why we ought always oppose overdetermination. For this to be a good reason, however, we need to be persuaded that overdetermination always involves redundancy and that the redundancy it involves is of an objectionable sort.

In the case of standard overdetermination, the redundancy of the overdetermining causes consists in the fact that each cause could have been deleted from the circumstances, nothing added in its place, without this affecting whether the effect came about. Such redundancy is objectionable, or at least it would be objectionable were it widespread and systematic, because it involves causes that are in a real sense surplus to requirement. Relating this to the passage from Merricks quoted at the start of this section, a substance dualist who accepted that certain physical effects always have both a physical cause and a mental cause that is a sort of *optional addition* to the physical cause, would be reasonably thought to be embracing redundancy of this kind.¹³ Let us assume that Merricks is right that, other things being equal, such a view ought to be rejected in favour of a view that does not need to embrace redundancy of this kind. That would not yet support his case against objects. For whatever the redundancy of objects might consist in, it cannot be of this form. While metaphysicians of objects disagree on whether or not particles arranged objectwise do in fact constitute objects, most of them agree – Merricks included – that *if* they do, they do so of metaphysical necessity. If that is right, then if there in fact *were* objects we simply *could not* delete them from the causal picture without thereby deleting the objectwise arrangements of particles that constitute them.

Similar considerations can be brought to bear on the mental case. Practically every contemporary philosopher of mind believes that the mental is in somehow intimately dependent on or determined by the physical. Minimally, they will accept some form of the supervenience thesis, which claims that no change in

¹¹ Of course there might be some specific account of causation that was incompatible with the occurrence of weak overdetermination. But if there is, we have not been told what this account is or why we ought to accept it. In section 6 below, I discuss whether a productive account of causation would be incompatible with weak overdetermination.

¹² A response to Sider's article (Merricks 2003) does little more than recap and refer back to this discussion.

¹³ I do not suggest that substance dualists need to hold this.

the mental is possible without some corresponding change in the physical. But in addition, many (perhaps most) of those who accept supervenience believe that it is to be explained in terms of dependence that involves not just nomological but metaphysical necessitation. There are different accounts of this dependence, but one idea is that the mental is ‘realized’ by the physical, where “[t]o realize is to ‘make real’ in a sense of ‘makes’ that is constitutive rather than causal” (Shoemaker 2007: 10). How one ought ultimately spell out the dependence is a question I will not attempt to answer here.¹⁴ It should be noted, however, that if the relationship between mental and physical entities is *constitutive*, then it is far from clear that overdetermination by mental and physical entities could involve redundancy of an objectionable form. For you *could not* delete mental causes from the causal picture without thereby deleting the physical entities that give rise to them.¹⁵

5. Parsimony

You might at this point appeal to parsimony. You might argue that a theory that has dependent causes is an ‘ugly theory’ and that if you can do without them you should. This seems at least sometimes to be what is motivating Merricks, and it may form part of Kim’s motivation too. So let me say a few words about why I don’t think appeal to parsimony would be successful.

Both Kim and Merricks appeal in their theories to the pluralities of physical particles and physical properties that play the causal roles of objects and mental properties on given occasions. Kim appeals to certain conjunctions of physical properties, whereas Merricks appeals to particles arranged objectwise. I submit further that if they are to stand a chance of giving a complete account of the causal picture, Kim and Merricks *must* appeal to such pluralities. It seems to me, however, that the need to appeal to such pluralities crucially undermines the claim to parsimony. First, it is not actually clear that there could be particles arranged tablewise, giving rise to all the properties and relations that we attribute to tables, *without* there being tables, because the conditions for there being tablewise arrangements of particles are more or less the same as the conditions for there being tables. This is not to say that the presence of the particles *per se* entails the presence of the table, nor even that the presence of the particles arranged in the form of a table entails the presence of a table. The claim is, rather, that the fact that the particles are *tablewise* arranged, which includes the assumption that all the properties that we attribute to tables are instantiated, entails that there is a table there. Given this, the claim that we ought to settle for particles arranged objectwise and abstain from objects brings to mind someone who upon being offered a slice of cake says ‘that’s very kind, thank you, but I am on a diet so I will make do with some cake particles arranged slice-wise.’

The theoretical savings are, in any case, questionable. For one thing, any theoretical savings that might be made in the number of things the theory postulates are likely to be matched by theoretical excesses incurred elsewhere. For instance, if one has a theory on which there are no objects or mental properties, one will need a much more complicated theory of meaning, to avoid the consequence that much of our talk is either false or nonsensical. And that is not all. Kim and Merricks need to make room for the

¹⁴ Shoemaker (2007) offers an account of realisation on which the causal powers of realised causes are a subset of the causal powers of their realisers. For an alternative, earlier version of the subset account see Jessica Wilson (1999, see also 2009).

¹⁵ Karen Bennett (2003) argues persuasively from such considerations that mental and physical causes cannot be taken to overdetermine their effects in the standard sense of overdetermination.

(purported) realizers of objects and mental properties in their theories. They do not merely need physical properties and particles. They need *certain pluralities* of physical properties and particles. More specifically, they need precisely those pluralities of physical properties and particles that we ordinarily think give rise to mental properties and objects. But then they owe us an account of how one might identify the relevant pluralities without appealing to the notion of objects and mental properties. If we have no criteria for what the relevant pluralities of particles and properties are that are not parasitic on our notions of objects and mental properties then there is no theoretical saving after all.

6. Causation as Production

In a recent paper, Barry Loewer (2007) suggests that if we think of causation in terms of production that would give us reason to accept the exclusion principle. Indeed, he submits, it is precisely because Kim is thinking of causation in these terms that he feels no need to argue for the principle.

My diagnosis of why Kim thinks that Exclusion is virtually analytic is that he is thinking of causation as *production*. If one thinks about causation in this way then it is quite natural to see Exclusion as virtually analytic. If $P(x,t)$ literally produces $Q(z,t')$ then it does appear that “there is no work” left for any other event $F(x,t)$ to do as far as producing $Q(z,t')$. Kim also seems to think of the causal relation as involving the transfer of some quantity, *causal oomph*, from the cause that brings the effect into existence. It is not surprising then that he would think that a second dose of *oomph* from $F(x,t)$ is not only not needed to produce $Q(z,t')$ there isn’t even any place for it. (Loewer, 2007, p.253)

I do not share Loewer’s belief that a productive account of causation would justify the exclusion principle. Nor am I convinced that Kim would want to *rely* on productionism to defend the principle. But I do think Kim would agree that the argument Loewer offers him here is a good one. (See the next section for an argument of Kim’s that is similar to Loewer’s offering, except that it makes no explicit appeal to productionism). And it is certainly true that in recent years Kim has openly embraced productionism. In a companion article to Loewer’s, Kim says:

An agent is someone who, because she has certain beliefs, desires, emotions, intentions, and the like, has the capacity to perform actions in the physical world—that is, to cause her limbs and other bodily parts (e.g., her vocal cords) move in appropriate ways so as to bring about changes in the natural course of events in the outside world—open a door, pick up the morning paper, and make a cup of coffee. It seems to me that without productive causation ... such causal processes are not possible. These causal processes all involve *real connectedness* between cause and effect, and the connection is constituted by phenomena like energy flow and momentum transfer, an actual movement of some (conserved) physical quantity. (Kim 2007, p.236)

Merricks too seems to be thinking of causation in terms of production. For instance, where particles arranged baseballwise make impact with a window, he claims that:

...each of the window-striking atoms causes *something*. And when you put what one atom causes together with what another causes, and so on for each of the atoms arranged baseballwise, it seems like the cumulative effect must be the shattering of the window. (Merricks 2001, p.62)

So it seems that both Kim and Merricks’s thinking is influenced by a productive view of causation. But would a productive account of causation really support the exclusion principle? And what would be the implications of this?

You might grant that if causation were correctly understood as production this would indeed vindicate the exclusion principle and rule out weak overdetermination, but then proceed to argue that believers in ordinary objects and irreducible mental properties need not be too upset by this, as we *need not* think of

causation in terms of production. Indeed, you might argue further that we *should not* think of causation in terms of production. This is Loewer's (2007) strategy. In the absence of other support for the exclusion principle, such a response would, if successful, show that the exclusion principle rests on an account of causation that is at best optional and at worst untenable. For what it's worth, I think such a response might well be successful. It is not, however, the response I would like to make. I do not think we need to argue against the productive view of causation in order to escape the exclusion problem. For I do not believe the productive view of causation would vindicate the exclusion principle. If anything, it seems to me that the exclusion principle looks *less* appealing if we accept a productive account of causation. For if we are thinking in terms of pushing and shoving, it seems that things that stand in the relation of constitution or realization would just be engaged, unproblematically, in the very same pushing and shoving on a given occasion.

Say I take a lump of clay, sculpt it into a statue, and then hurl it at a window with the result that the window breaks. Many metaphysicians want to say that the statue and the lump are numerically distinct.¹⁶ But, on the current suggestion, a productive account of causation will not allow this. Echoing Loewer's passage, if the lump's hitting the window literally produces the window breaking, then there is no work left over for the statue's hitting the window to do in producing the window breaking. More, not only is a second dose of oomph from the statue not needed to produce the window breaking, there is no place for it. All of this looks true to me. What I fail to see, however, is why this should make us believe that the statue, conceived as a thing numerically distinct from but constituted by the lump, could not have been involved in the window breaking. The natural thing to say here, it seems to me, is that the statue and the lump on this occasion imparted the very same causal powers on the window. If we cannot say this, we have not yet been told why.

What the production account might give us is one way of making plausible that realized or constituted causes do not on a given occasion have any additional causal powers to the causes that realize or constitute them. It does nothing, however, to show that the realized or constituted causes would *need* to have such additional causal powers. And that, as I shall now argue, is the claim that Kim and Merricks's arguments most crucially rest on.

7. The Principle of Additional Causal Powers

The moral that I would draw from the above considerations is that where two causes are as intimately related as objects and their particles or mental properties and the physical properties that realize them, the sufficiency of one cause does not entail the redundancy of the other. Interestingly, however, both Kim and Merricks believe that, far from providing dissolution of the problem, the dependence relationship between the purported causes is partly what *generates* the problem. Here is Kim:

The exclusion problem doesn't go away when we recognize the two purported causes as in some way related to each other, perhaps one being dependent on the other. As long as they are recognized as distinct events, each claiming to be a full cause of a single event, the problem remains. (1998: 53)

¹⁶ This is because the statue and the lump of clay appear to have different properties. E.g. the lump existed before the statue did, and the lump could survive squashing whereas the statue could not.

It is important to see that the problem that we face arises *because* the two putative causes are *not* independent events. The difficulty is exactly that the causal status of the dependent event is threatened by the event on which it depends. (1998: 53).

But why exactly should we think that the causal status of the dependent cause threatened by the cause on which it depends? Kim's reason, it turns out, is that where one purported cause is intimately dependent on another, the dependent cause cannot bring any *additional causal powers* to those of the cause on which it depends. And this, he evidently thinks, means that the purported dependent cause couldn't really be a cause after all.

[T]here is a real problem, the exclusion problem, in recognizing second-order properties as causally efficacious *in addition to* their realizers (1998, p.53, my emphasis).

For there is nothing in the instantiation of F on this occasion over and above the instantiation of its realizer H. Given this, to think that this instance of F has causal powers *in excess* of these of H is tantamount to belief in magic. (1998, p.54-55, my emphasis)¹⁷

I think we should grant Kim that mental property instantiations do not, *on any given occasion*, have causal powers in excess of those of the property instantiations that realize them. More generally, I think we should grant that where one entity is realized or constituted by another, the constituted or realized entity could not, *on any given occasion*, bring with it causal powers in excess of those already brought by the entities that constitute it.¹⁸ I think we should deny, however, that this 'failure' to bring additional causal powers precludes dependent entities from being causes.

This is where we reach the bedrock of Kim's argument, and indeed of my disagreement with him. It boils down to the following principle, which Kim accepts and I reject:¹⁹

The Principle of Additional Causal Powers: Where an effect has a sufficient cause C1, any numerically distinct and simultaneously occurring sufficient cause C2 would have to bring with it causal powers additional to those had by C1 in the given circumstances.

If this principle were true, it seems it would not only make Kim's exclusion argument successful. It would also make Merricks's argument successful. For baseballs could not, according to this principle, be sufficient causes of window breakings, as baseballs do not, on any given occasion, bring with them causal powers that are *additional* to those of the particles arranged baseballwise on that occasion.²⁰ Indeed, I think Merricks himself implicitly appeals to something like this principle. It is hinted at, for instance, in passages such as the following:

Obviously enough, one would have a reason for believing that the shattering of the window is overdetermined only if one had a reason for believing that *both* the baseball *and* the atoms arranged baseballwise caused it (Merricks 2001: 72 original emphases).

More, Merricks evidently thinks that one could have such a reason only if the baseball had causal powers *in addition to* those of the particles arranged baseballwise. He offers the following analogy:

¹⁷ The sentiment in these passages is similar to that expressed by Loewer in the preceding section, though Kim does not (explicitly at least) rely here on any particular account of causation. It should be noted also that although the passages quoted appear in the context of a discussion of the exclusion problem as it presents itself for functionalism, Kim's line of thought is not restricted to some particular version of functionalism. Indeed, Kim himself stresses that the exclusion problem "is entirely general; it arises for all second-order properties, whether or not they are functional." (1998: 55).

¹⁸ Please see section 8 for further discussion of this point.

¹⁹ I have elsewhere, in a paper that contains precursors to some of the ideas developed here, referred to this as 'the motivating principle' (see Árnadóttir and Crane 2013).

²⁰ I am granting Merricks here that particles arranged baseballwise have the causal powers we attribute to baseballs. For further thoughts on this see section 8.

Imagine that someone has been killed by a bullet. Now entertain the possibility that the killing was overdetermined by two bullets arriving simultaneously. But suppose, further, that there is *no reason* to believe that the killing was overdetermined in this way. For, let us suppose, while there is evidence for the existence of one bullet, there is no evidence for the existence of a second. In such a case, I think everyone would agree that we should deny that the killing is overdetermined as a result of a second bullet. For without a reason to think an effect is overdetermined, we should assume it is not. (2001: 72)

The implication, I take it, is that while there is evidence for the existence of particles arranged baseballwise, there is no evidence for the existence of a baseball, given that the particles alone can account for all the causal work that we might attribute to the baseball, including the causal impact on our senses. These considerations are epistemological and are aimed at undermining the perceptual evidence that we take ourselves to have for the existence of objects. But the principle of additional causal powers is lurking in the background. The thought seems to be that if the object has no causal powers for impinging upon our senses *additional* to those of the particles arranged objectwise, then there could be no reason to believe in them. Something similar seems to be motivating Kim when he says:

[I]f the causal powers of the M-instance are identical with those of its realizing P-instance, what reasons can we have for thinking that there are two events, not one? (Kim 1998: 56).

The intended moral, I take it, is that if the purported second cause brings no additional causal powers to the first, then *there could be no reason* to think that it genuinely exists.

It seems to me then that both Kim and Merricks's arguments at least implicitly rely on the principle of additional causal powers. It seems to me further that both Kim and Merricks *require* this principle. Without it, we have no reason to accept either Kim's exclusion principle or Merricks's claim that objects do not overdetermine their effects.

8. *Having Causal Powers vs. Putting them in Place*

In the next section, I shall argue that we ought to reject the principle of additional causal powers. But first I want briefly to discuss a worry that one might have concerning the constituting or realizing entities *having* the causal powers attributed to the constituted or realized entities rather than merely *putting them in place*.

Anyone who believes that mental causes are realized by physical causes will agree that the causal powers of mental causes are put in place in any given circumstances by the physical causes that realize them. And anyone who believes that objects are constituted by particles arranged objectwise will agree that the causal powers of objects are put in place in any given circumstances by the particles that constitute them. So in that sense, at least, everyone should agree that mental causes and ordinary objects cannot bring any additional causal powers to the scene. I.e. they cannot bring any causal powers that are not already put in place by their realizers or constituting entities. The way I formulated Kim and Merricks's arguments above, this was the only concession they needed. But you might worry that the principle of additional causal powers requires more than this. It requires not only that the constituting causes put all the relevant causal powers in place, but that the constituting causes *themselves have* all these causal powers. And that, you might think, is not obviously a concession that the believers in irreducible mental causes and ordinary objects need to make. Given this, it is not clear that the principle of additional causal powers would in fact validate the arguments.

Although I think it worth noting that this line of response is possible, I will not here attempt to develop it. Doing so would require going into much more detail about different accounts of causation than I am willing or able to do here, and it would throw me off the course that I want to take. My aim is not to establish how exactly we should think of mental causation or object causation, but rather to show that Kim and Merricks's arguments fail on their own terms. This I hope to do by identifying the most basic assumption on which their arguments rest and showing why we need not grant that assumption, even granting the other claims required for the arguments' success. I am, in any case, inclined to agree with Kim and Merricks that particles arranged objectwise and physical property instantiations do, on any given occasion, have the causal powers that they give rise to. For it seems to me very reasonable to think that one way of having causal powers on a given occasion is to constitute or realize something that has them.²¹ In light of this, I will grant in the ensuing discussion that the constituting or realizing causes have the causal powers of the would-be constituted or realized entities.

I should stress that if it turned out that I was wrong about this, if it could be argued that the constituting and/or realizing causes did not in fact have all the causal powers we attribute to the would-be constituted and/or realized entities, that would not undermine my case against the principle of additional causal powers. Rather, it would give us an alternative case against Kim and/or Merricks's use of the principle.²² It would allow us to argue that the principle fails to validate the argument(s). In order to validate their arguments, Kim and/or Merricks would then require the following strengthened version of the principle:

The Strengthened Principle of Additional Causal Powers: Where an effect has a sufficient cause C1, any numerically distinct and simultaneously occurring sufficient cause C2 would have to bring with it causal powers additional to those put in place by C1 in the given circumstances.

But, I submit, Kim and Merricks's opponents are even less likely to accept this version of the principle than the version stated above. (To verify this, you are invited to plug the strengthened version into the arguments of the following section).

These qualifications out of the way, I turn now to explaining why we cannot employ the principle of additional causal powers to make good Kim and Merricks's arguments, even assuming that the constituting or realizing causes have all the causal powers we attribute to the constituted or realized causes.

9. Rejecting the Principle of Additional Causal Powers

Let me start by acknowledging that I think the principle of additional causal powers has some surface plausibility. Indeed I have found, in discussions with various colleagues, that many are inclined to accept it on gut feeling, although I have yet to be offered an argument in its favour. This gut reaction may partly explain why Kim's argument has generated such a large variety of responses that grant the exclusion

²¹ This is perhaps most obviously true if we think of causation in terms of production, though I suspect it will be true on other accounts of causation too, as long as we keep in the picture both the intimate relationship that holds between the causes, and the restriction of the causal claims to given occasions. But my inclination to agree with Kim and Merricks is not based on any particular account of causation, but rather on the naïve thought that what a thing (or things) can do depends on what it makes up.

²² I say 'Kim and/or Merricks's use of the principle' for one might attempt here to drive a wedge between their arguments, e.g. by arguing that while physical causes have the causal powers we attribute to mental causes, particles arranged objectwise do not have the causal powers we attribute to objects, though they do put them in place. This is what Kim himself wants to say (see his 2005, p.56).

principle but challenge the argument on other grounds.²³ Gut reactions aside, however, I believe that closer inspection shows that we ought to reject the principle.

I believe the principle of additional causal powers derives its plausibility from the plausibility of a similar but importantly distinct principle known as Alexander's dictum, which claims that 'to be is to have causal powers'. As we saw at the outset, versions of this principle are central to both Kim and Merricks's work. Kim claims that in order to exist, mental properties must have causal powers, whereas Merricks claims that in order to exist, ordinary objects must have causal powers. Both of these claims are, moreover, intuitively plausible. The idea of an ordinary object that is absolutely devoid of causal powers is at best mysterious, and the idea of a property that is absolutely devoid of causal powers is no less mysterious. Indeed it might even be thought to be incoherent. I submit that one reason that people are attracted to the principle of additional causal powers is that they don't properly distinguish it from Alexander's dictum. It is very important, however, to distinguish the two principles. And it is very important to note that Alexander's dictum does not entail the principle of additional causal powers.

While Alexander's dictum demands only that a thing have *some* causal powers, the principle of additional causal powers demands that a thing have *additional* causal powers. In slogan form, we might put this as saying that "two things cause more than one" or that "to be is to have *additional* causal powers". This principle is stronger than Alexander's dictum and is not entailed by it. For a thing might well have causal powers without having *additional* causal powers, in the sense intended. It is important here also that the principle of additional causal powers is meant to apply to *particular circumstances*. It claims that where an effect has a sufficient cause C1, any numerically distinct and simultaneously occurring sufficient cause C2 would have to bring with it causal powers additional to those of C1 *in the given circumstances*. When Kim and Merricks demand additional causal powers, they are demanding that the entity in question have some additional causal powers *on a given occasion*. They are focussing in on singular causal statements, or the states of affairs that such statements describe, and demanding that the purported cause have some causal powers that are not *on that occasion* had by anything distinct from it. But I see no good reason to think that lacking such additional causal powers would strip a thing of its causal powers. (For some reasons to think it would not, see section 10 below). If Kim and Merricks know of such reasons, they do not make them explicit.

I would argue further that it only seems plausible that lacking such additional causal powers would strip a thing of its causal powers if we ignore the possibility that the purported causes might be dependent. I would like to draw an analogy at this point, that I hope might be instructive, between Kim and Merricks's reasoning and David Lewis's 'weight argument' against coincident objects. Lewis says:

It reeks of double counting to say that here we have a dishpan, and we also have a dishpan-shaped bit of plastic that is just where the dishpan is, weighs just what the dishpan weighs (why don't the two together weigh twice as much?), and so on. This multiplication of entities is absurd on its face. (Lewis, 1986: 252).

A central thought here is that if the dishpan and the dishpan-shaped bit of plastic were distinct things, then the dishpan ought to have some *additional* weight to that of the plastic. In a slogan "two material objects

²³ While most philosophers of mind appear to agree that there is *something* wrong with Kim's argument, people's responses differ widely. Some have concluded that we must return to some form of reductive physicalism (e.g. Papineau 1990; Kim 1998, 2005), while others have thought that the solution lies, rather, in adopting some form of dualism (e.g. Lowe 2008). And a number of authors have suggested that we solve the problem by focusing either on the relation of causation (e.g. Robb 1997; Ehring 1999; Macdonald and Macdonald 1986; Gibb 2004) or on the causal relation (e.g. Loewer 2007; Menzies 2008; Raatikainen 2010).

weigh more than one” or even perhaps “to be a material object is to have additional weight”. The combined weight of the dishpan and its plastic, however, is exactly the same as the weight of the plastic. So, the thought runs, we ought to conclude that there is no such thing as a dishpan that is distinct from the plastic. Although this argument gets some approving nods in the literature, it is cheap. Anyone who believes in distinct but coincident objects will distinguish between two kinds of non-identity; *distinctness*, which requires only that the things in question are not numerically identical, and *separateness*, which requires minimally also that the things in question be composed of different matter. The problem with the weight principle is that it is plausible only when restricted to separate things. If the dishpan and plastic were separate things, then we would of course expect their combined weight to be double that of one. But they are *not* separate things. They share all their matter. And given this, we would expect their combined weight to be exactly the same as the weight of one of them. Of course, there may well be good reasons for denying that there are distinct coincident objects. But the fact that each adds nothing to the weight of the other is not one of them. Indeed, arguing that coincident objects would *have to* add to each other’s weight in order to exist straightforwardly begs the question against those who claim that there are such things.

It seems to me that the principle of additional causal powers, together with the evidential considerations in its favour, similarly begs the question against those who claim that there are objects and irreducible mental causes. Take the view that there are such things as baseballs and that baseballs are constituted by, but not identical to, particles arranged baseballwise. It will not do to argue against this view by complaining that a baseball, *on a given occasion*, has no causal powers in addition to those of the particles arranged baseballwise. For if the particles arranged baseballwise did in fact compose a baseball, this is precisely what you would expect. More generally, it will not do to argue against the view that particles arranged objectwise compose objects by complaining that an object does not, on a given occasion, have any causal powers in addition to those of the particles that compose it. For if particles arranged objectwise did in fact compose objects, then this is precisely what you would expect. In a sense – though *not* in the sense of strict numerical identity, and therein lies the gap – the baseball just *is* its constituting particles on that occasion. Given this, it seems that the baseball could not, on a given occasion, have causal powers that are not had by its physical realizers.²⁴

Similar considerations apply to the mental case. If mental causes are constituted or realized by physical causes, then it will not do to argue that mental causes must be identical to, rather than realized by, physical causes on grounds that mental causes don’t have any causal powers in excess of those of their realizers on any given occasion. For if mental causes were in fact realized by, rather than identical to, physical causes, then that is precisely what we should expect. Nothing more is required for the presence of the mental cause, on a given occasion, than the presence of the physical entities that realize it on that occasion. For instance, nothing more is required for the instantiation of a mental property on a given occasion than the instantiation of the physical properties that realize it on that occasion. Given this, it seems that the mental property could not, on any given occasion, have causal powers that are not had by its physical realizers.

²⁴ As explained above, I am assuming here that the particles’ arranged baseballwise giving rise to the would-be causal powers of the baseball on a given occasion amounts to their having these causal powers on that occasion. If this were shown to be false, Merricks would need to retract to the strengthened version of the principle of additional causal powers, and the strengthened version would certainly be shown to be untenable by the considerations just given.

The moral is that the additional causal powers principle is plausible only when restricted to separate and independent causes. The same can be said for the evidential considerations offered by Merricks. In the case of the two bullets, what makes it reasonable to request extra causal evidence for the existence of the second bullet and to doubt the existence of the second bullet in the absence of such evidence is that the two purported causes are not merely numerically distinct, but also separate and independent entities. Where the purported causes stand in some relation of realization or constitution to one another, things look very different. For given the nature of such relationships we could not expect to find, when homing in on any given occasion, any causal evidence for the constituted entity beyond the causal evidence that we have for the entities that constitute it. I do not know whether there are any good reasons for denying that there are objects and mental properties. But if there are, the fact that objects and mental properties do not on any given occasion cause anything in addition to their realizers is not one of them. Indeed, claiming that objects and mental properties would have to have additional causal powers of this kind, or that we could have no reason to believe in them unless they did, in the absence of some independent arguments for these claims, is just to beg the question against those who claim that there are objects and mental properties.

10. Challenges for Believers in Objects and Mental Properties

Insofar as my task was to show that Kim and Merricks's arguments fail and to explain how they fail, I have little more to say. Before concluding, however, I would like to say a few words about certain challenges that the arguments might be thought to raise for believers in objects and mental properties. For you might think that even if they are ultimately unsuccessful, the arguments nonetheless challenge the believers in objects and mental properties to explain how these things could play any causal role at all, and what reasons we could have for believing in them, given that they have no additional causal powers to those of their realizers on any given occasion. These are big challenges and I cannot hope to address them satisfactorily here. I would like to say a few words, however, that I hope will at least show that the onus is on the objectors to explain why we ought to think these challenges *cannot* be met.

The suspicion that these challenges present a serious problem arises partly, I think, from a metaphysical approach that proposes to construct the world from snapshots of microphysical reality at moments in time. If we home in on particular causal situations, whether we are focusing on particular instantiations of properties or on objects and particles at particular moments in time, we are liable to lose sight of any causal differences between things that stand in the relation of realization or constitution. Indeed, as far as singular causal statements involving particular circumstances are concerned, it seems there are no causal differences. For there is nothing more to the realized entity *on the given occasion* than the realizing entity or entities. That is not to say, however, that there might not be any causal differences to be gleaned when we zoom back out. Mental properties, for instance, trace out causal explanatory patterns that are not traced out by any physical properties. It is not just that mental properties can be realized by different physical properties on different occasions, but also that the instantiation of any mental property on a given occasion is realized by the instantiation of a myriad of physical properties. So even as far as particular instantiations of mental properties are concerned, we could not hope to identify the mental property with any neat physical property. At best, we could hope to identify such instantiations with some *plurality* of physical property instantiations. Typically, mental properties are said to be irreducible because there are too many physical properties that can play their roles. But if we focus on mental properties as *types* or *universals*,

rather than as *instantiations* thereof, the problem is not really that there are *too many* candidate physical properties for mental properties to be reduced to, but that there are *none*. For there are in fact *no* physical properties that play the exact roles of mental properties across the various circumstances in which they are instantiated. Given this, if we were to reduce mental properties (types) we would first have to *produce* some physical properties (types) to reduce them to. Arguably, the moral of this is that mental properties, conceived of as types or universals, have causal powers not had by any neat physical properties, conceived of as types or universals.²⁵ That is one reason to think that an entity might have an important causal role to play even though it brings no additional causal powers to its realizer(s) on a given occasion.

Here is another reason. One could coherently hold that certain pluralities of physical properties on a given occasion have certain of their causal powers *by virtue of* realizing the relevant mental property. Put simply, the thought would be that the physical has certain causal powers precisely *because* it gives rise to the mental. One need not adopt any full-blown emergentist account to make sense of this thought. Just suppose that the very same conditions must be met for the physical realizers to have certain causal powers and for them to realize a given mental property. It seems plausible that there might be such conditions. And it seems reasonable to think that where such conditions are met this would be a case of some plurality of physical property instantiations simultaneously having certain causal powers and realizing a mental property instance that has these same causal powers. Similar considerations apply to the objects case. One might hold that the very same conditions must be met for particles to have or give rise to some causal powers and for them to compose a certain sort of object. Again, it seems plausible that there are such conditions. And it seems reasonable to think that where they are met this would be a case of particles simultaneously having certain causal powers and composing an object that has these same causal powers. Further, it seems natural to think that this might be a case of particles having certain causal powers precisely *by virtue of* composing an object with these causal powers. A lot more would need to be said, of course, but in rough outline, these ideas look coherent. If they are, then an entity may have a distinctive and important causal role to play even though it adds nothing, on a given occasion, to the causal powers of its realizers.

How about the second question? What reasons could we have for believing in objects and mental properties? I do not think it incumbent on the believer in objects to answer this question. Pre-theoretically, it seems we have every reason to believe in objects and mental properties. If we were to accept that we do not in fact have such reasons, then we would need to be presented with some very strong theoretical defeaters of our reasons. The burden of proof is very much on those who argue that there are no objects and/or mental properties. Kim believes that the fact that mental properties do not have additional causal powers to those of their realizers on a given occasion seriously undermines our reasons to believe that there are mental properties distinct from these realizers. Similarly, Merricks believes that the fact that objects do not have additional causal powers to those of their would-be constituting particles on a given occasion seriously undermines our reasons to believe that there are any objects. But Kim and Merricks still owe us arguments for why we ought to accept these claims. In the absence of such arguments, the principle of additional causal powers and the evidential considerations in its favour beg the question against believers in objects and mental properties.

²⁵ I hasten to add that if any form of physicalism is true, then these causal powers will of course be *determined* by physical properties.

11. Concluding Remarks

In assessing overdetermination arguments, we must keep firmly in mind the distinction between a thing's making a causal difference and a thing's making an *additional* causal difference to that of the phenomena that give rise to it on a given occasion. When you throw a baseball at a window it does make a causal difference. But it does not make an *additional* causal difference to that exerted by the particles arranged baseballwise *in these particular circumstances*. In his book, *Physicalism or something near enough*, Kim says: "If the argument goes wrong, one would like to know just where and how it goes wrong." It will need to be shown, he continues "that there are these causal relations *in addition to* the underlying physical causal processes" (Kim, 2005, p.54). I think this last sentence is instructive about where both Kim and Merricks's arguments go wrong. They go wrong precisely in assuming that if *a* is not identical to *b*, then one must always *add something* to a circumstance in which you already have *b* in order to get *a*. They go wrong in assuming that if *a* is not identical to *b*, then *a* must be something *over and above* *b*, *on that particular occasion*. On a metaphysical picture on which things stand in the realization or constitution to one another this assumption is not warranted. For on such a picture, there are ways of being *nothing over and above*, on a given occasion, that do not involve identity. If the believers in objects and irreducible mental causes are right, then for the same reason that objects and mental causes would not add anything to the picture, they also do not need adding to the picture. They are already there. Indeed, they are necessarily there. Once this is appreciated, the threat of exclusion and elimination of both ordinary objects and causally efficacious mental entities evaporates.²⁶

References:

- Árnadóttir, S.T. and Crane, T. (2013). 'There is No Exclusion Problem', in S. Gibb and R. Ingthorsson (eds.), *Mental Causation and Ontology*. Oxford University Press: 248-266.
- Bennett, K. (2003). 'Why the exclusion problem seems intractable, and how, just maybe, to tract it', *Nous* 37: 471-497.
- Ehring, D. (1999). 'Tropelless in Seattle: The Cure for Insomnia', *Analysis* 59(261): 19-24.
- Fodor, J. (1974). 'Special sciences', *Synthese* 28: 97-115.
- Gibb, S.C. (2004). 'The Problem of Mental Causation and the Nature of Properties', *Australasian Journal of Philosophy* 82(3): 464-476.
- Kim, J. (1998). *Mind in a Physical World: An Essay on the Mind-Body Problem and Mental Causation*. Cambridge, MA: Bradford.
- Kim, J. (2005). *Physicalism or Something Near Enough*. Princeton: Princeton University Press.
- Kim, J. (2007). 'Causation and Mental Causation', in B.P. McLaughlin and J.D. Cohen (eds.), *Contemporary Debates in Philosophy of Mind*. Oxford: Blackwell: 227-242.
- Lewis, D. (1986). *On the Plurality of Worlds*. Oxford: Basil Blackwell.
- Loewer, B. (2007). 'Mental Causation or Something Near Enough', in B.P. McLaughlin and J.D. Cohen (eds.), *Contemporary Debates in Philosophy of Mind*. Oxford: Blackwell: 243-264.
- Lowe, E.J. (2008). *Personal Agency: The Metaphysics of Mind and Action*. Oxford: Oxford University Press.
- Macdonald, C. and G. Macdonald (1986). 'Mental Causes and Explanation of Action', *Philosophical Quarterly* 36: 145-158.
- Menzies, P. (2008). 'The Exclusion Problem, the Determination Relation and Contrastive Causation', in J. Hohwy and J. Kallestrup (eds.), *Being Reduced: New Essays on Reduction, Explanation, and Causation*. New York: Oxford University Press: 196-217.
- Merricks, T. (2001). *Objects and Persons*. Oxford: Clarendon Press.
- Merricks, T. (2003). 'Replies', *Philosophy and Phenomenological Research* 67(3): 727-744.
- Papineau, D. (1990). 'Why Supervenience?', *Analysis* 50(2): 66-71.
- Papineau, D. (2002). *Thinking About Consciousness*. Oxford: Oxford University Press.
- Raatikainen, P. (2010). 'Causation, Exclusion, and the Special Sciences', *Erkenntnis* 73: 349-363.
- Robb, D. (1997). 'The Properties of Mental Causation', *Philosophical Quarterly* 47: 178-194.
- Shoemaker, S. (2007). *Physical Realization*. Oxford: Oxford University Press.

²⁶ The first draft of this paper was prepared while holding a Leverhulme Early Career fellowship. Thanks to Tim Crane, John Donaldson, two reviewers for this journal, and audiences at London, Cambridge and Manchester for comments on earlier versions.

- Sider, T. (2003). 'What's so Bad About Overdetermination?', *Philosophy and Phenomenological Research* 67(3): 719-726.
- van Inwagen, P. (1990). *Material Beings*. Ithaca: Cornell University Press.
- Wilson, J. (1999). 'How Superduper does a Physicalist Supervenience Need to Be?', *The Philosophical Quarterly* 49: 33-52.
- Wilson, J. (2009). 'Determination, realization and mental causation', *Philosophical Studies* 145: 149-169.